

Fully Adaptive Ridge Detection Based on STFT Phase Information

Marcelo A. Colominas[✉], Sylvain Meignen[✉], and Duong-Hung Pham[✉]

Abstract—This letter deals with the problem of the estimation of the instantaneous frequencies of the modes of multicomponent signals from their linear time-frequency representations. In most approaches, such an estimation consists of extracting the ridges associated with each mode in the time-frequency plane. A major issue associated with these techniques is that ridge detection relies on some ad-hoc parameters which essentially bound the modulation of the studied modes and put some constraints on the type of filter used in the time-frequency representation. In this paper, we alternatively propose a novel fully adaptive approach for ridge detection whose relevance is shown throughout numerical simulations.

Index Terms—Time-frequency, short-time Fourier transform, synchrosqueezing, instantaneous frequency, ridge detection.

I. INTRODUCTION

MULTICOMPONENT signals, defined as superimpositions of amplitude- and frequency-modulated (AM-FM) waves, have been an active subject of study for at least 50 years. Such signals are encountered in various practical non-linear systems, for instance, pathology diagnosis [1]–[3], or structural damage [4], [5]. The most common way to deal with them is to compute their linear time-frequency (TF) representation such as the short-time Fourier transform (STFT) or the continuous wavelet transform (CWT), and then look for the *ridges* of its different *modes* in the TF plane. These ridges constitute estimations of the instantaneous frequencies (IFs), which are often used subsequently for the retrieval of the modes [6]. Alternatively, there exist many other techniques performing mode reconstruction without considering TF representation (TFR), as for instance *variational methods* [7], [8], *empirical wavelet transform* [9] or *empirical mode decomposition* [10]. However, all of them basically extract band-limited modes, and thus are essentially different from those obtained with TF techniques.

Traditionally, ridge detection (RD) is carried out on the *modulus* of the TFR, but this way to proceed ignores half of the information of the latter, namely its *phase* [11], [12]. Our goal

is thus to show that this phase information, which is often estimated using reassignment operators in *reassignment techniques* like *synchrosqueezing transforms* (SSTs) [13]–[18], is useful to design adaptive RD. Indeed, a serious issue related to RD from TFR is that it often relies on some arbitrary upper bound on the modulation of the modes and depends on some hypotheses made on the filter used for TFR computation. Therefore, we propose, in this paper, a fully adaptive RD computed on STFT magnitude using modulation operators introduced in [15], [17]. To do so, we first recall, in Section II, some useful definitions and notations, and then a commonly used parametric RD based on STFT. Our main contribution, in Section III, then explains how to use the just mentioned modulation operators to design a more adaptive RD. Finally, numerical simulations carried out in Section IV validate the new technique on both artificial and real signals.

II. MULTICOMPONENT SIGNALS, STFT, RIDGE DETECTION AND MODE RECONSTRUCTION

Multicomponent signals (MCSs) consist of the superimposition of a finite number L of components modulated both in amplitude and frequency (AM-FM):

$$x(t) = \sum_{l=1}^L x_l(t) = \sum_{l=1}^L A_l(t) e^{i2\pi\phi_l(t)}, \quad (1)$$

where $A_p(t)$ and $\phi'_p(t)$ are, respectively, the instantaneous amplitude (IA) and instantaneous frequency (IF): $A_l(t), \phi'_l(t) > 0 \forall t$, $\phi'_k(t) - \phi'_l(t) > 0 \forall k > l$ and $|A'_l(t)|$ are assumed to be small. The (modified) Short-Time Fourier Transform (STFT) corresponds to [19]:

$$F_x^g(t, f) = \int_{-\infty}^{+\infty} x(u)g(u-t)e^{-i2\pi f(u-t)}du, \quad (2)$$

where $g(t)$ is an even real window with $\text{supp}\{\hat{g}(f)\} \subseteq [-B, +B]$.¹ In a discrete-time setting, $x[n]$ is a finite time series of length N , defined on $[0, 1]$, and its STFT is written as:

$$F_x^g\left(\frac{n}{N}, \frac{kN}{K}\right) \approx F_x^g[n, k] := \sum_{u=0}^{N-1} x[u]g[u-n]e^{-i2\pi\frac{k(u-n)}{K}},$$

with $k \in \{0, \dots, K-1\}$ and $n \in \{0, \dots, N-1\}$.

When a signal is modeled as in (1), every component occupies a “ribbon” around its IF $\phi'_l(t)$. In addition, if the IFs are separated

Manuscript received February 20, 2020; revised March 31, 2020; accepted April 6, 2020. Date of publication April 14, 2020; date of current version May 7, 2020. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jens Ahrens. (Corresponding author: Marcelo A. Colominas.)

Marcelo A. Colominas is with the Institute for Research and Development in Bioengineering and Bioinformatics (IBB), CONICET, Entre Ríos, Argentina (e-mail: macolominas@conicet.gov.ar).

Sylvain Meignen is with the Jean Kuntzmann Laboratory, University of Grenoble-Alpes and CNRS UMR 5224, 38041 Grenoble, France (e-mail: sylvain.meignen@univ-grenoble-alpes.fr).

Duong-Hung Pham is with the IRIT Laboratory, CNRS UMR 5505, 31400 Toulouse, France (e-mail: duong-hung.pham@irit.fr).

Digital Object Identifier 10.1109/LSP.2020.2987166

¹ $\hat{g}(f) = \int_{\mathbb{R}} g(t)e^{-i2\pi ft}dt$ is the Fourier transform (FT) of $g(t)$.

enough (i.e. $|\phi'_k(t) - \phi'_l(t)| > 2B, \forall k \neq l$), then each mode “lives” in non-overlapping domains [11], [15]. Besides, RD consists of extracting the dominant lines, called *ridges* around which the ribbons are arranged, by solving iteratively the following optimization problem, for l from 1 to L by means of a peeling algorithm [12], [20]:

$$\max_{c_l} \sum_{n=0}^{N-1} |F_{x,l}^g[n, c_l[n]]|^2, \quad \text{s.t. } |\Delta c_l[n]| \frac{N^2}{K} \leq B_f, \quad (3)$$

where c_l is the l^{th} estimated ridge, $\Delta z[n] = z[n+1] - z[n]$, while B_f is a *fixed a priori* upper bound for the frequency modulation of the modes. Note that L is assumed to be known, and we refer to [21] for an analysis on how to assess the number of ridges. The constraint in (3) means that $c_l[n+1] \in [c_l[n] - \frac{B_f K}{N^2}, c_l[n] + \frac{B_f K}{N^2}]$. Further, $F_{x,l}^g$ is recursively defined as follows:

$$F_{x,l}^g[n, k] = \begin{cases} 0, & \text{if } k \in [c_l[n] - \eta_l^-[n], c_l[n] + \eta_l^+[n]] \\ F_{x,l-1}^g[n, k], & \text{otherwise,} \end{cases} \quad (4)$$

where $F_{x,1}^g = F_x^g$ (i.e. RD for the first ridge is carried out on the initial STFT), while η_l^- and η_l^+ should be defined. It is important to emphasize that the quality of RD strongly depends on how the latter are defined. For example, the most traditional RD technique sets them to constant values: $\eta_l^-[n] = \eta_l^+[n] = \lceil \frac{BK}{N} \rceil$, where $\lceil \cdot \rceil$ denotes the *ceiling* function, and B is the half-width of $\text{supp}\{\hat{g}\}$. In what follows, this technique is denoted by CFB-RD (CFB for *constant frequency bandwidth*).

Alternatively, one can make η_l^- and η_l^+ time-dependent by exploring the noise behavior of the signal in the TF plane. To this end, one considers a noisy signal \tilde{x} modeled by $\tilde{x}[n] = x[n] + r[n]$, where r is a complex white Gaussian noise with variance σ_r^2 . The STFT of the noise, being also Gaussian with zero mean, satisfies: $\text{var}(\Re\{F_r^g[n, k]\}) = \text{var}(\Im\{F_r^g[n, k]\}) = \sigma_r^2 \|g\|_2^2$, where $\Re\{X\}$ and $\Im\{X\}$ stand for the real and imaginary parts of complex number X , respectively, and $\|\cdot\|_2$ for the ℓ_2 norm. Also, it was shown in [22] that $|F_r^g[n, k]|^2 / \sigma_r^2 \|g\|_2^2$ is χ^2 distributed with 2 degrees of freedom, and a threshold of 9 corresponds to a probability of false alarm smaller than 1%. Therefore, the thresholded STFT with threshold $T = 3\sigma_r \|g\|_2$ can be written as:

$$\bar{F}_{\tilde{x}}^g[n, k] = \begin{cases} F_{\tilde{x}}^g[n, k], & \text{if } |F_{\tilde{x}}^g[n, k]| \geq T = 3\sigma_r \|g\|_2 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

Note that in practical situations, $\sigma_r \|g\|_2$ is not known, and can be estimated, for example, by $\text{median}(|\Re\{\bar{F}_{\tilde{x}}^g[n, k]\}|) / 0.6745$ [23], [24]. Inspired by (5), one proposes to define $\eta_l^-[n]$ and $\eta_l^+[n]$ such that $[c_l[n] - \eta_l^-[n], c_l[n] + \eta_l^+[n]]$ is the longest interval containing $c_l[n]$ in which the modulus of the STFT remains larger than T [24], [25]. In the sequel, such a strategy is denoted by VFB-RD (VFB for *varying frequency bandwidth*).

III. A NEW STRATEGY FOR RD

The above introduced RD techniques depend greatly on an a priori fixed parameter B_f , and on how the bandwidth functions

η_l^- and η_l^+ are defined. Indeed, on one hand, to fix B_f a priori somewhat limits the adaptivity of RD, and, on the other hand, regarding the definition of bandwidth functions, CFB-RD does not take into account potential frequency modulation of the modes. Finally, VFB-RD requires an estimation of the noise level, and is thus inappropriate in noiseless situations. Therefore, we propose herein a novel technique that adaptively modifies RD by first getting rid of B_f and then by defining new bandwidth functions.

A. Frequency Modulation Operators

Let us first introduce an estimate of the frequency modulation of the modes used in the definition of *second order STFT-based synchrosqueezing transform* (FSST2) [15], which relies on a complex estimate:

$$\tilde{q}_{\tilde{x}}[n, k] = \frac{1}{2i\pi} \frac{F_{\tilde{x}}^{g''}[n, k] F_{\tilde{x}}^g[n, k] - (F_{\tilde{x}}^{g'}[n, k])^2}{F_{\tilde{x}}^{tg}[n, k] F_{\tilde{x}}^{g'}[n, k] - F_{\tilde{x}}^{tg'}[n, k] F_{\tilde{x}}^g[n, k]}, \quad (6)$$

in which $F_{\tilde{x}}^{g'}$, $F_{\tilde{x}}^{tg}$, $F_{\tilde{x}}^{g''}$, $F_{\tilde{x}}^{tg'}$ are respectively STFTs of f computed with windows $n \mapsto g'[n]$, $(tg)[n]$, $g''[n]$ and $(tg')[n]$ is such that $\hat{q}_{\tilde{x}}[n, k] := \Re\{\tilde{q}_{\tilde{x}}[n, k]\}$ consists of an approximation of $\phi''(\frac{n}{N})$, when x is a linear chirp. In the case of an MCS, when $k \frac{N}{K}$ lies in the vicinity of $\phi'_l(\frac{n}{N})$, $\hat{q}_{\tilde{x}}[n, k]$ is an estimate of $\phi'_l(\frac{n}{N})$ [17]. In what follows, we are going to explain how to use $\hat{q}_{\tilde{x}}[n, k]$ in the definition of a new adaptive RD.

B. RD Based on Frequency Modulation Estimate

As highlighted above, a great limitation associated with existing RD is that the upper bound B_f is kept *constant* regardless of the frequency modulation of the modes. So, using the frequency modulation estimate $\hat{q}_{\tilde{x}}[n, k]$ described in Section III-A, we now propose to improve RD by relaxing the constraint on B_f in the optimization problem (3) as follows:

$$\begin{aligned} \max_{c_l} \sum_{n=0}^{N-1} |F_{\tilde{x},l}^g[n, c_l[n]]|^2, \\ \text{s.t. } \left| \Delta c_l[n] - \frac{K}{N^2} \hat{q}_{\tilde{x}}[n, c_l[n]] \right| \leq C, \end{aligned} \quad (7)$$

where C is some integer corresponding to the number of frequency bins $\Delta c_l[n]$ can depart from $\frac{K}{N^2} \hat{q}_{\tilde{x}}[n, c_l[n]]$. It is worth mentioning that in (3), no assumption were made on the sign of the modulation because B_f was a bound to the absolute value of the maximum modulation, while in (7), c_l and $\Re\{\tilde{q}_{\tilde{x}}[n, c_l[n]]\}$ are constrained to have the same sign. Moreover, unlike (3), the constraint in (7) means that $c_l[n+1] \in [c_l[n] + \frac{K}{N^2} \hat{q}_{\tilde{x}}[n, c_l[n]] - C, c_l[n] + \frac{K}{N^2} \hat{q}_{\tilde{x}}[n, c_l[n]] + C]$, with the interval center being a “predictor” for the ridge at time $n+1$.

After having relaxed the constraint on B_f , we now define new frequency bandwidth functions, $\eta_l^-[n]$ and $\eta_l^+[n]$, associated with the ridge c_l at each time instant n . For that purpose, let us consider the l^{th} mode: $x_l(t) = A_l(t) e^{i2\pi\phi_l(t)}$, which *locally* admits a linear chirp approximation [15], [26]: $x_l(u) \approx A_l(t) e^{i2\pi(\phi_l(t) + \phi'_l(t)(u-t) + \frac{\phi''_l(t)}{2}(u-t)^2)}$, for u sufficiently close

to t . Its STFT is then approximated as:

$$F_{x_i}^g(t, f) \approx x_i(t) \widehat{g_{\phi_i''}(f - \phi_i'(t))}, \quad (8)$$

where $g_{\phi_i''}(t) = g(t)e^{i2\pi\frac{\phi_i''}{2}t^2}$. It is interesting to note that contrary to CFB-RD which ignores the frequency modulation term $e^{i2\pi\frac{\phi_i''}{2}t^2}$, we propose to use it in the computation of new bandwidth functions. It was shown in [26] that for a Gaussian window $g(\frac{n}{N}) = e^{-\sigma(\frac{n}{N})^2}$, and for a given time instant n , we have that $|F_{x_i}^g[n, k]|$ is almost zero when $|\frac{kN}{K} - \phi_i'(\frac{n}{N})| \geq \frac{3}{\pi} \sqrt{\frac{\sigma^2 + \pi^2 \phi_i''(\frac{n}{N})^2}{2\sigma}}$. Using also the frequency modulation estimate $\hat{q}_x[n, c_i[n]]$ of $\phi_i''(\frac{n}{N})$, as introduced in Section III-A, we define: $\eta_i^-[n] = \eta_i^+[n] = \left\lfloor \frac{K}{N} \frac{3}{\pi} \sqrt{\frac{\sigma^2 + \pi^2 \hat{q}_x^2[n, c_i[n]]}{2\sigma}} \right\rfloor$, as frequency bandwidth functions when the signal is noisy. The just introduced RD is called MB-RD (MB for *modulation-based*) in the sequel. Note that this technique could be easily extended to TFRs based on CWT since the modulation operators are also computable in that context [27].

IV. EXPERIMENTS AND RESULTS

This section presents some numerical experiments conducted on artificial and real signals to illustrate the improvement achieved by the proposed approach MB-RD over the state-of-the-art methods, CFB-RD and VFB-RD.

A. Artificial Signals

First of all, we consider synthesized signals which are uniformly sampled over time interval $[0, 1]$ with sampling rate at $M = 1024$ Hz. Their STFTs are computed with $K = N$ frequency bins and with a Gaussian window whose length σ is determined by the minimum of the Rényi entropy computed on the STFT magnitude [28]. The signals are also contaminated by white Gaussian noises with different *signal-to-noise ratios* (SNRs). The performance of RD is evaluated by $SNR_{out}(\phi_e') = 20 \log_{10}(\|\phi_{ref}'\|_2 / (\|\phi_e' - \phi_{ref}'\|_2))$, where ϕ_e' is the estimated IF corresponding to the detected ridges and ϕ_{ref}' is the reference IF.

1) *Monocomponent Signals*: Let us first consider monocomponent artificial signals $x_p(t) = e^{i2\pi\phi_p(t)}$, $p = 1, 2$, in which the frequency modulation is important. Since in the monocomponent case CFB-RD and VFB-RD are the same, only CFB-RD is reported here. In particular, we investigate the following two cases:

- $\phi_1(t) = 250t + 25e^{-120(t-\frac{1}{2})^2}$, with $\phi_{1,ref}'(t) = 250 - 6000(t - \frac{1}{2})e^{-120(t-\frac{1}{2})^2}$.
- $\phi_2(t) = 50t + \int (\frac{N}{5}\text{erf}(25(t-\frac{1}{2}))) + \frac{N}{5} dt$, with $\phi_{2,ref}'(t) = 50 + \frac{N}{5}\text{erf}(25(t-\frac{1}{2})) + \frac{N}{5}$, where $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the error function.

On these signals, MB-RD is compared with CFB-RD computed with B_f equal either to $2N$, $5N$ or $50N$. On the first row of Fig. 1, we depict the moduli of their STFTs, while the performance of RD with these techniques are displayed on the second row. For the first signal (*first column*), it can be shown that $|\max(\phi_1''(t))| = 6000$, and therefore, as expected, CFB-RD with $B_f = 2N$ performs poorly. On the contrary, when B_f is set to a bigger value, for instance $B_f = 50N$, such a technique

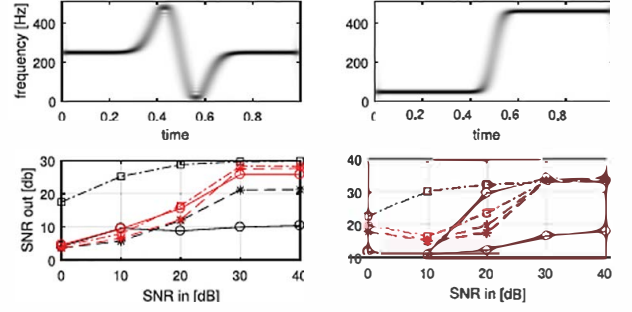


Fig. 1. First row: STFT moduli of the monocomponent signals. Second row: CFB-RD (black) using $B_f = 2N$ (circles), $B_f = 5N$ (stars), or $B_f = 50N$ (squares). MB-RD (red) using $C = 2$ (circles), $C = 5$ (stars), or $C = 10$ (squares). The average performance over 30 realizations is displayed.

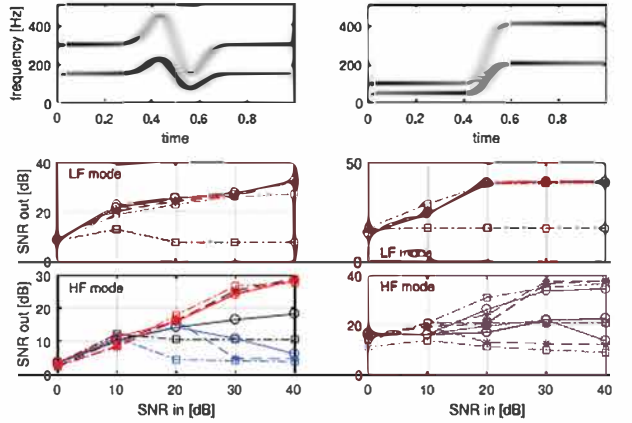


Fig. 2. First row: STFT moduli of the multicomponent signals. Second and third rows: CFB-RD (black) and VFB-RD (blue) using $B_f = 2N$ (circles), $B_f = 5N$ (stars), or $B_f = 50N$ (squares). MB-RD (red) using $C = 2$ (circles), $C = 5$ (stars), or $C = 10$ (squares). The average performance over 30 realizations is displayed.

behaves nicely. On the contrary, MB-RD appear to be slightly sensitive to the choice of C , behaves similarly to CFB-RD with $B_f = 5N$, and performs consistently better as the noise level decreases (input SNR of 30 or 40 dB). For the second signal (*second column*), one has $|\max(\phi_2''(t))| = 5N$, and thus CFB-RD with $B_f = 2N$ behaves poorly, while MB-RD remains only relatively insensitive to C . In addition, the RD results given by MB-RD or CFB-RD with $B_f = 5N$ are still of comparable quality, but clearly MB-RD has the advantage to be fully adaptive. In short, when using CFB-RD on a monocomponent signal, B_f can be chosen arbitrarily large so that it is always more efficient than MB-RD. However, we are going to see, in the following subsection, that this is no longer the case when dealing with MCSs.

2) *Multicomponent Signals*: We now investigate synthetic signals made of two components, both of which are amplitude-modulated and the second component has its phase equal to twice that of the first one: $x(t) = a(t)e^{i2\pi\phi_p(t)} + b(t)e^{i2\pi2\phi_p(t)}$, $p = 1, 2$, with $a(t) = 1 + e^{-10(t-1/2)^2}$ and $b(t) = 2 - e^{-10(t-1/2)^2}$. To be in line with the study of the monocomponent case, we consider the following phase functions:

- $\phi_1(t) = 150t + 8e^{-120(t-\frac{1}{2})^2} 100t^2$, with $\phi_{1,ref}'(t) = 150 - 1920(t - \frac{1}{2})e^{-120(t-\frac{1}{2})^2}$.

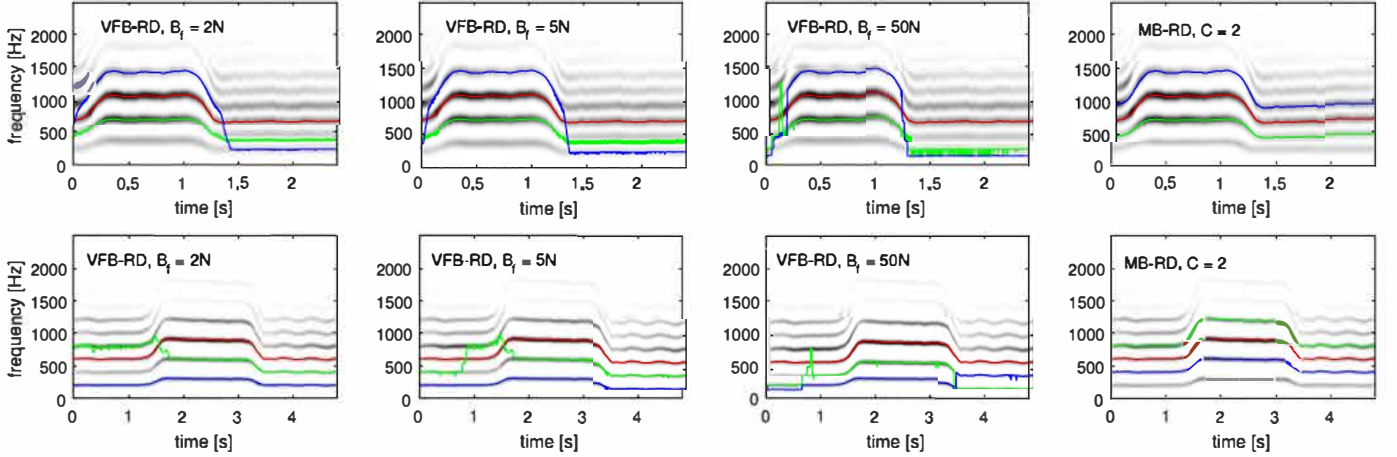


Fig. 3. RD results of the different studied methods on: (first row), recording 8; (second row) recording 62 [29]. Red: first detected ridge. Green: second detected ridge. Blue: third detected ridge. In these simulations, we use $K = N/4$ for the computation of the STFT.

- $\phi_2(t) = 50t + \int (\frac{3N}{40} \text{erf}(20(t - \frac{1}{2})) + \frac{3N}{40}) dt$, with $\phi'_{2,\text{ref}}(t) = 50 + \frac{3N}{40} \text{erf}(20(t - \frac{1}{2})) + \frac{3N}{40}$.

The STFT moduli of such signals are displayed on the first row of Fig. 2, while the second row exhibits the performance of RD with the different methods. It should be noted that the *lower frequency* (LF) mode is always detected first therefore the behaviors of CFB-RD and VFB-RD are the same for that mode. Thus, on the second row of Fig. 2, we only display the results related to CFB-RD, along with those corresponding to MB-RD. We first remark that these two techniques produce similar results on the LF mode provided B_f is not chosen too large in CFB-RD, namely if mode mixing does not drive down the performance, as for example when $B_f = 50N$. Unlike in the monocomponent case, a large value for B_f is no longer allowed, since the ridge detector might then “jump” from one ridge to another at two successive time instants. On the contrary, MB-RD exhibits little dependence on the value of C .

Then, since RD methods are all associated with a peeling scheme, after having extracted the first ridge, the neighborhood of this ridge is removed in the TF plane before proceeding with the extraction of the second one. It should be noted that this neighborhood depends on the RD technique used: with VFB-RD, at high input SNRs, the frequency bandwidth parameters $\eta_1^-[n]$ and $\eta_1^+[n]$ (corresponding here to the LF mode) are too large which is problematic when the modes exhibit interference in the TF plane (around $t = 0.5$ in the studied cases). This results in a failure of VFB-RD on the *high frequency* (HF) mode (third row of Fig. 2). Switching to the results calculated by CFB-RD, we notice that, as expected, for the HF mode, this technique is more sensitive to B_f than for the LF mode (see third row of Fig. 2). On the contrary, MR-RD behaves similarly for all these modes, with only a slight dependency on parameter C .

B. Real Signals

Voice signals have been widely studied using TF analysis techniques [30]–[33], and we here investigate how the proposed MB-RD technique can be beneficial in that context. To do so,

we consider two signals from the Saarbruecken voice database [29], both of which are sustained /a/ vowels of the type “low-high-low”, meaning there is a change on their pitch.

The first signal corresponds to the recording of a 19-years-old healthy female speaker (called recording 8 [29]). On the first row of Fig. 3, we depict its STFT magnitude, along with the RD results given by VFB-RD with $B_f = 2N$, $5N$ or $50N$ (first to third columns) and those computed with MB-RD (fourth column). It is clear that only MB-RD accurately extracts the three ridges, while VFB-RD leads to ridges containing “jumps” and undesired fast oscillations, regardless of the value of B_f . Moreover, on the second row of Fig. 3, we display the same computations as above, but for the signal corresponding to a recording of a 22-years-old healthy female speaker (recording 62 [29]). Note that such a signal is different from the previous one in the fact that it exhibits a *vibrato* on its final third. We clearly observe from these figures that VFB-RD leads to mode mixing whatever the value of B_f : the algorithm “jumps” from one ridge to another. On the contrary, MB-RD catches perfectly the *vibrato* on the final third pitch of the signal.

V. CONCLUSIONS

In this letter, a novel fully adaptive technique for the detection of time-frequency ridges associated with the modes of a multicomponent signal was proposed. It was based on the local estimate of the frequency modulation of the modes used for the computation of the second-order synchrosqueezing transform. In particular, a new unsupervised way to dealing with the jump when detecting ridges was introduced, which showed the importance of taking into account the frequency modulation when using a peeling algorithm. Numerical experiments demonstrated the effectiveness of the proposed technique on both simulated and real signals. Future work will be dedicated to embed this new technique into the downsampled short-time Fourier transform [25] or synchrosqueezing transform-based demodulation context [34], for the purpose of mode reconstruction.

REFERENCES

- [1] S. Cerutti, A. L. Goldberger, and Y. Yamamoto, "Recent advances in heart rate variability signal processing and interpretation," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 1, pp. 1–3, Feb. 2006.
- [2] U. R. Acharya, K. P. Joseph, N. Kannathal, L. C. Min, and J. S. Suri, "Heart rate variability," in *Advances in Cardiac Signal Processing*. Berlin, Germany: Springer, 2007, pp. 121–165.
- [3] M. Malik and A. J. Camm, *Dynamic Electrocardiography*. Hoboken, NJ, USA: Wiley, 2008.
- [4] O. Salawu, "Detection of structural damage through changes in frequency: A review," *Eng. Struct.*, vol. 19, no. 9, pp. 718–723, 1997.
- [5] C. R. Farrar, S. W. Doebling, and D. A. Nix, "Vibration-based structural damage identification," *Philos. Trans. Royal Soc. London A: Math., Physical Eng. Sci.*, vol. 359, no. 1778, pp. 131–149, 2001.
- [6] F. Auger *et al.*, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013.
- [7] S. Chen, X. Dong, Z. Peng, W. Zhang, and G. Meng, "Nonlinear chirp mode decomposition: A variational method," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 6024–6037, Nov. 2017.
- [8] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2013.
- [9] J. Gilles, "Empirical wavelet transform," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 3999–4010, Aug. 2013.
- [10] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. Royal Soc. London A: Math., Physical Eng. Sci.*, vol. 454, pp. 903–995, 1998.
- [11] R. Carmona, W. L. Hwang, B. Torrésani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Trans. Signal Process.*, vol. 45, no. 10, pp. 2586–2590, Oct. 1997.
- [12] R. Carmona, W. L. Hwang, B. Torrésani, "Multiridge detection and time-frequency reconstruction," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 480–492, Feb. 1999.
- [13] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," *Wavelets Medicine Biol.*, pp. 527–546, 1996.
- [14] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 243–261, 2011.
- [15] T. Oberlin, S. Meignen, and V. Perrier, "Second-order synchrosqueezing transform or invertible reassignment? Towards ideal time-frequency representations," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1335–1344, Mar. 2015.
- [16] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis - with an application to gravitational-wave signal," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 3168–3178, Jun. 2017.
- [17] R. Behera, S. Meignen, and T. Oberlin, "Theoretical analysis of the second-order synchrosqueezing transform," *Appl. Comput. Harmon. Anal.*, vol. 45, no. 2, pp. 379–404, 2018.
- [18] H. Yang, "Statistical analysis of synchrosqueezed transforms," *Appl. Comput. Harmon. Anal.*, vol. 45, no. 3, pp. 526–550, 2018.
- [19] L. Cohen, *Time-Frequency Analysis*, vol. 1, Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.
- [20] D. Iatsenko, P. V. McClintock, and A. Stefanovska, "Extraction of instantaneous frequencies from ridges in time-frequency representations of signals," *Signal Process.*, vol. 125, pp. 290–303, 2016.
- [21] N. Saulig, N. Pustelnik, P. Borgnat, P. Flandrin, and V. Sucic, "Instantaneous counting of components in nonstationary signals," in *Proc. 21st Eur. Signal Process. Conf.*, 2013, pp. 1–5.
- [22] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. New York, NY, USA: Academic, 2009.
- [23] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [24] D.-H. Pham and S. Meignen, "A novel thresholding technique for the denoising of multicomponent signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2018, pp. 4004–4008.
- [25] S. Meignen and D.-H. Pham, "Retrieval of the modes of multicomponent signals from downsampled short-time Fourier transform," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6204–6215, Dec. 2018.
- [26] M. A. Colominas, S. Meignen, and D.-H. Pham, "Time-frequency filtering based on model fitting in the time-frequency plane," *IEEE Signal Process. Lett.*, vol. 26, no. 5, pp. 660–664, May 2019.
- [27] T. Oberlin and S. Meignen, "The second-order wavelet synchrosqueezing transform," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2017, pp. 3994–3998.
- [28] R. G. Baraniuk, P. Flandrin, A. J. Janssen, and O. J. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.
- [29] B. Woldert-Jokisz, "Saarbrücken voice database," 2007.
- [30] R. K. Potter, G. A. Kopp, and H. C. Green, *Visible speech*, 1947.
- [31] L. Pimonow, *Vibrations en Régime Transitoire*. Paris, France: Dunod, 1962.
- [32] M. Sondhi and J. Schroeter, "A hybrid time-frequency domain articulatory speech synthesizer," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 7, pp. 955–967, Jul. 1987.
- [33] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1830–1847, Jul. 2004.
- [34] D.-H. Pham and S. Meignen, "Demodulation algorithm based on higher order synchrosqueezing," in *Proc. 27th Eur. Signal Process. Conf.*, 2019, pp. 1–5.